

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 02-232697

(43)Date of publication of application : 14.09.1990

(51)Int.Cl.

G10L 3/02

(21)Application number : 01-053200

(71)Applicant : NIPPONDENSO CO LTD

(22)Date of filing : 06.03.1989

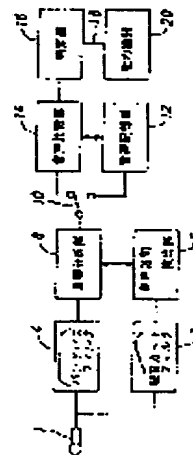
(72)Inventor : SHIMIZU KIMIHARU
YAMAOKA HARUYASU
SUZUKI KUNIKAZU
NAKAMURA KAZUO
UONO YUTAKA
ASADA HIROSHIGE

(54) VOICE RECOGNITION DEVICE

(57)Abstract:

PURPOSE: To improve the recognition rate by providing an acoustic analyzing means which extracts feature data from a voice signal corresponding to a voice section of the voice signal from a microphone and outputs it to a comparing means.

CONSTITUTION: A noise cut filter 2 outputs the voice signal inputted from the microphone 1 after attenuating its band where more noise components are present and a voice section detection means 6 detects a voice section with the signal from the noise cut filter 2. Then an acoustic analyzing means 8 extracts the feature data from the voice signal corresponding to the voice section of the voice signal from the microphone 1 and outputs the data to the comparing means 14, which compares standard data stored in a voice storage means 12 with the feature data. Consequently, the feature data can be extracted from the inputted voice signal not through the noise cut filter 2 and the voice can be recognized based upon the feature data, so the recognition rate is improved.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision
of rejection]

[Date of requesting appeal against examiner's
decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

⑩ 日本国特許庁(JP)

⑪ 特許出願公開

⑫ 公開特許公報(A)

平2-232697

⑬ Int.Cl.⁵

G 10 L 3/02

識別記号

3 0 1

庁内整理番号

8842-5D

⑭ 公開 平成2年(1990)9月14日

審査請求 未請求 請求項の数 2 (全6頁)

⑮ 発明の名称 音声認識装置

⑯ 特 願 平1-53200

⑰ 出 願 平1(1989)3月6日

⑱ 発 明 者	清 水	公 治	愛知県刈谷市昭和町1丁目1番地	日本電装株式会社内
⑱ 発 明 者	山 岡	晴 康	愛知県刈谷市昭和町1丁目1番地	日本電装株式会社内
⑱ 発 明 者	鈴 木	邦 一	愛知県刈谷市昭和町1丁目1番地	日本電装株式会社内
⑱ 発 明 者	中 村	一 雄	愛知県刈谷市昭和町1丁目1番地	日本電装株式会社内
⑱ 発 明 者	宇 尾	野 豊	愛知県刈谷市昭和町1丁目1番地	日本電装株式会社内
⑱ 発 明 者	浅 田	博 重	愛知県刈谷市昭和町1丁目1番地	日本電装株式会社内
⑲ 出 願 人	日本電装株式会社			愛知県刈谷市昭和町1丁目1番地
⑳ 代 理 人	弁理士 足 立 勉			外2名

明 細 書

1 発明の名称

音声認識装置

2 特許請求の範囲

(1) 音声に応じた標準データを予め記憶する音声記憶手段を有し、前記標準データとマイクロフォンから入力される音声信号に応じた特徴データとを比較する比較手段を備えた音声認識装置において、

前記音声信号の内の騒音成分を多く含む帯域を減衰して出力する騒音カットフィルタと、

該騒音カットフィルタからの信号により音声区間を検出する音声区間検出手段と、

前記マイクロフォンからの音声信号の内、前記音声区間に応じた前記音声信号から特徴データを抽出して前記比較手段に出力する音声分析手段と、を備えたことを特徴とする音声認識装置。

(2) 前記フィルタは、人間の聴感特性に合致あるいは近似の補正特性を持つフィルタであることを特徴とする特許請求の範囲第1項記載の音声

認識装置。

3 発明の詳細な説明

[産業上の利用分野]

本発明は、音声情報を判定する音声認識装置に関し、特に、非定常騒音の多い環境下における音声認識装置に関する。

[従来の技術]

従来より、キーワードとなる操作者の発声する音声信号と、登録済みの音声信号との類似度により、音声信号を認識して、各種機器の運転を制御する音声認識装置が知られている。そして、音声認識装置の用いられる周囲の環境からの騒音によって、例えば、空気調和装置に音声認識装置を組み込んだ場合に、空気調和装置の持つ騒音、振動音、及びその他の外部の騒音によって、音声の誤認識を起こす場合があり、その対策が取られている。

例えば、空気調和装置の運転中には、音声信号を第5図に示すような低周波帯域をカットする特性を有する騒音カットフィルタを通すことによっ

て、騒音の信号を減衰させている。その結果、第6図に示すような低周波数の空気調和装置の発生する騒音あるいは外部の騒音を第7図に示すように減衰することができる。そして、騒音カットフィルタを通した音声信号から特徴データを抽出し、予め記憶された標準データとを比較して、そのマッチングによって音声情報を判定し、空気調和装置等を制御していた。

〔発明が解決しようとする課題〕

しかしながら、こうした従来の音声認識装置では、入力される音声信号を騒音カットフィルタにより処理するので、騒音を低減することはできるが、騒音の低減と共に音声信号の一部までも減衰してしまう場合があった。例えば、音声信号として、母音「あ」が入力された場合には、第8図に示すような周波数成分の音声信号がマイクロフォンから出力される。そして、この音声信号を前述した特性を有する騒音カットフィルタにより処理すると、音声の特徴付けるホルマント周波数の内、低周波数の第1、第2ホルマント等を第8図

間を検出する音声区間検出手段と、

前記マイクロフォンからの音声信号の内、前記音声区間に応じた前記音声信号から特徴データを抽出して前記比較手段に出力する音響分析手段と、
を備えたことを特徴とする音声認識装置の構成がそれである。

〔作用〕

前記構成を有する音声認識装置は、騒音カットフィルタが、マイクロフォンから入力された音声信号の内の騒音成分を多く含む帯域を減衰して出力し、音声区間検出手段が、騒音カットフィルタからの信号により音声区間を検出する。そして、音響分析手段が、マイクロフォンからの音声信号の内、音声区間に応じた音声信号から特徴データを抽出して比較手段に出力し、比較手段が音声記憶手段に記憶された標準データと前記特徴データとを比較する。よって、騒音カットフィルタを通すことなく入力された音声信号から特徴データを抽出でき、この特徴データに基づいて音声の認識ができるので認識率が向上する。

に示す斜線の如く減衰させてしまい、そのスペクトルのピークを欠落させてしまう。その為、フィルタ処理後の特徴抽出の際に、音声信号の内、認識時最も重要となる特徴量を減少させることとなってしまう、認識率の低下を招くという問題があった。

そこで本発明は上記の課題を解決することを目的とし、騒音を減衰させるフィルタの影響を受けることなく音声信号の特徴を抽出し、認識率の向上を図った音声認識装置を提供することにある。

〔課題を解決するための手段〕

かかる目的を達成すべく、本発明は課題を解決するための手段として次の構成を取った。即ち、

音声に応じた標準データを予め記憶する音声記憶手段を有し、前記標準データとマイクロフォンから入力される音声信号に応じた特徴データとを比較する比較手段を備えた音声認識装置において、

前記音声信号の内の騒音成分を多く含む帯域を減衰して出力する騒音カットフィルタと、

該騒音カットフィルタからの信号により音声区

〔実施例〕

以下本発明の実施例を図面に基づいて詳細に説明する。

第1図は本発明の一実施例である音声認識装置の概略構成図である。1は周知のマイクロフォンであり、操作者の発した音声を変換し音声信号として出力するものである。このマイクロフォン1は、騒音カットフィルタ2と、バンドパスフィルタ4とにそれぞれ接続されている。騒音カットフィルタ2は、マイクロフォン1から入力される騒音を含んだ音声信号から騒音成分を含んでいる周波数帯域を減衰させるものである。本実施例では、第2図に示すような低周波数と高周波数の領域の音声を大きく減衰させ2kHz～4kHz付近で最も感度が高い、人間の聴感特性に合致した補正特性を有するものである。

あるいは、このような聴感補正特性を有するものの以外でも、その環境により、騒音のスペクトル成分に高い周波数成分が多い場合には、騒音カットフィルタとしてハイカットフィルタを用いれば

よい。又、第6図に示すような低周波成分を多く含む騒音特性を有する環境下におかれている場合には、その環境に応じて第5図に示すような低周波成分を減衰させる周波数特性を有するものであってもよい。若しくは、騒音が特定の中域に集中している場合には、騒音カットフィルタとして中域カットフィルタを用いればよい。更には、プログラムによってその特性を変えることができる自由度の大きなデジタルフィルタを用い、周囲の騒音状態等によって、その特性を切り換える構成としてもよい。尚、フィルタの減衰率等の特性を強く設定し過ぎると、騒音はより減衰させられるが、同時に純粋な音声信号も減衰させられてしまう。又、特性を弱くすると騒音の減衰が弱く、後述する音声区間の検出に影響がでる。

そして、もう一方のバンドパスフィルタ4は、マイクロフォン1から入力される騒音を含んだ音声信号を、音声帯域、例えばおおそ2.00Hz～4KHzの周波数帯域以外の成分をカットするものである。このフィルタとして、デジタルフ

ィルタを用いてもよい。

前記騒音カットフィルタ2は、音声区間検出部6に接続されており、この音声区間検出部6は、騒音カットフィルタ2から出力された音声信号のパワー情報と予め設定された所定のしきい値とを比較して、音声区間の始端と終端を検出し、しきい値のレベルより高いときは音声区間として、低いときは無音区間として出力するものである。このしきい値は、固定した1つだけでもよいが、複数のしきい値や周囲の騒音に合わせてしきい値を可変にしてもよい。また、音声スペクトルの傾きやピッチ情報などと併せて、音声区間を検出するものであってもよい。この音声区間検出部6は、前記バンドパスフィルタ4と共に音響分析部8に接続されており、音響分析部8には、バンドパスフィルタ4を通った音声帯域の音声信号と、音声区間信号とが入力される。尚、騒音カットフィルタ2を通過した信号によって音声区間を検出するだけでなく、この信号に基づいて図示しないアンプのゲインをコントロールするようにしてもよい。

入力信号が大きければアンプのゲインを下げ、小さければアンプのゲインを上げることによって、音声信号のダイナミックレンジを大きくすることができる。この時、基準となる信号に騒音成分が多く含まれていると正確なゲインコントロールは不可能であり、騒音カットフィルタ2により騒音成分をカットした信号を基準にすることにより、より正確なゲインコントロールが可能となる。

この音響分析部8は、入力された音声区間信号に基づいて、その音声区間内のバンドパスフィルタ4を通った音声帯域の音声信号のスペクトルを分析し、音声の特徴を表す特徴パラメータを抽出するものである。この抽出に当たっては、例えば周知の高速フーリエ変換(FFT)、バンドパスフィルタバンクや、線形予測分析などにより一定周期毎に、入力された音声区間内の音声信号から特徴パラメータを抽出して、特徴ベクトルの時系列として出力するものである。尚、音響分析部8において、更に、騒音成分を低減するような処理を行って音声信号の分析をするようにしてもよい。

この音響分析部8は、切換スイッチ10を介して、音声記憶部12若しくは音声比較部14に選択的に接続できるようになされている。前記音声記憶部12は、抽出された特徴データ、例えば、ベクトルの時系列を標準データとして記憶するものである。そして、前記音声比較部14は、この音声記憶部12に記憶された標準データと、切換スイッチ10を介して入力される特徴データとを比較し、その類似度を計算して出力するものである。この音声比較部14は、判定部16に接続されており、判定部16は、音声比較部14からの類似度に応じてそれと最も似通ったかつ予め決められた一定の条件を満たした場合に、それに対応した信号を出力するものである。この信号は出力端子18から出力されるようになされており、この出力端子18に接続された他の機器20、例えば空気調和装置を制御するようになされている。

次に、本実施例の作動について説明する。

まず、操作者が音声を発声すると、例えば「うんてん」と発声すると、それがマイクロフォン1

によって拾われて、電気信号に変換されて音声信号として出力される。この時、マイクロフォン1によって、発声された音声の前後での音声以外の外部からの騒音も入力され、第3図に示すような騒音が重畳された音声信号が出力される。

この出力される音声信号は、それぞれ騒音カットフィルタ2とバンドパスフィルタ4とに入力される。騒音カットフィルタ2に入力された音声信号は、第4図に示す如く音声信号に重畳された騒音成分を減衰させられ音声区間検出部6に出力される。音声区間検出部6では、入力された音声信号のパワー情報と所定のしきい値とを比較して、このしきい値より低い区間は無音区間と判定し、しきい値よりも高い区間は音声区間と判定し、音声区間信号を音響分析部8に出力する。尚、騒音カットフィルタ2に人間の聴感補正特性に合致した特性をもたせ、人間の感覚に似た分析を行って、理想的な音声認識の分析とすることができる。

一方、バンドパスフィルタ4は、入力された音声信号を音声帯域以外を減衰させて音響分析部8

される。そして、音声比較部14では、特徴データと標準データとの類似度を計算して出力する。判定部16では、この音声比較部14から出力される類似度に基づいて操作者が発した音声があるかを判断し、その発声した音声に応じた信号を出力する。そして、出力端子18を介して、その信号を他の機器20に出力する。

例えば、空気調和装置に用いた場合には、その発声が「うんてん」である場合には、空気調和装置の運転を開始する信号を出力する。あるいはその発声に応じた信号を出力して、設定温度を上げたりする制御を実行する。

尚、本実施例では、特定話者の音声認識を例としたが、不特定話者の音声認識であっても同様に実施可能である。

前述した如く、本実施例の音声認識装置は、騒音カットフィルタ2により騒音を減衰し、音声区間検出部6がその音声信号に基づいて音声区間を検出する。そして、この音声区間と騒音カットフィルタ2を通らない音声信号とに基づいて音響分

に出力する。音響分析部8では、このバンドパスフィルタ4からの出力された音声信号と、音声区間検出部6からの音声区間信号とから、音声区間内の音声信号を分析して特徴データを抽出する。即ち、操作者によって音声を発声された区間の音声信号のみに基づいて特徴データを抽出し、それ以外の区間は、騒音であるとして特徴データを抽出する処理を行わない。

そして、切換スイッチ10が音声記憶部12側に切り換えられているときには、前記音響分析部8から出力される特徴データを標準データとして音声記憶部12に記憶する。こうして、音声記憶部12には、特定話者が発声した単語や短文等の音声の特徴データ、例えば、音声が「うんてん」である場合にはその特徴データが標準データとして記憶される。

切換スイッチ10が音声比較部14側に切り換えられているときには、前記音響分析部8から出力される特徴データと、前記音声記憶部12に記憶された標準データとが、音声比較部14で比較

析部8が特徴データを抽出する。切換スイッチ10を切り換えて、この特徴データを標準データとして音声記憶部12に記憶し、又、音声比較部14によって、特徴データと標準データとを比較する。

従って、音響分析部8においては、騒音カットフィルタ2を通らない音声信号から特徴データを抽出する。よって、騒音カットフィルタ2により音声成分までも減衰された音声信号からの音声の特徴データの抽出ではなく、減衰されていない音声信号から適正に特徴データを抽出することができる。これにより、的確に音声の特徴を掴むことができ、音声比較部14で特徴データと標準データとの類似度の計算精度が向上し、音声の認識率が向上する。

又、騒音が発声音声にオーバーラップしていても、騒音カットフィルタにより騒音成分を減衰させた信号から音声区間検出を実行することで、精度よく音声区間検出を行うことが可能である。従来は発声音声の音声区間に騒音が重なって、音声区間

が広く取られ、これによって、特徴データを抽出していたので、認識エラーを引き起こす場合があったが、適正に音声区間を検出することによって、認識率の向上を図ることができる。

以上本発明はこの様な実施例に何等限定されるものではなく、本発明の要旨を逸脱しない範囲において種々なる態様で実施し得る。

〔発明の効果〕

以上詳述したように本発明の音声認識装置は、騒音成分を含んだ入力音声から騒音成分を減衰させ、その信号から音声区間検出を実行することで、精度よく音声区間検出を行うことが可能であり、かつ音響分析部に音声情報を何ら欠落させることなく入力して分析することを可能とし、騒音が大きい環境、特に非定常的な騒音の多い環境での認識性能向上にすぐれた効果がある。

4 図面の簡単な説明

第1図は本発明の一実施例としての音声認識装置の概略構成図、第2図は聴感補正特性を有する騒音カットフィルタの特性を示すグラフ、第3図

はマイクロフォンからの音声信号を示すグラフ、第4図は騒音カットフィルタにより処理した信号のグラフ、第5図はローカットフィルタの特性を示すグラフ、第6図は低域に集中した騒音のレベルを示すグラフ、第7図はローカットフィルタにより第6図の騒音を処理したレベルのグラフ、第8図はホルマント周波数を説明するグラフである。

1…マイクロフォン

2…騒音カットフィルタ

6…音声区間検出部

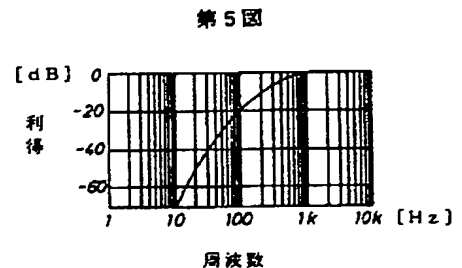
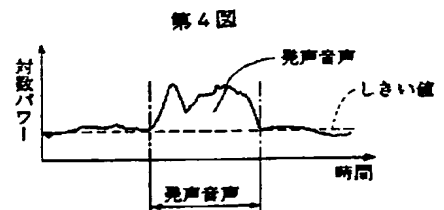
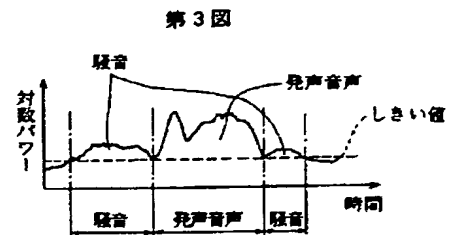
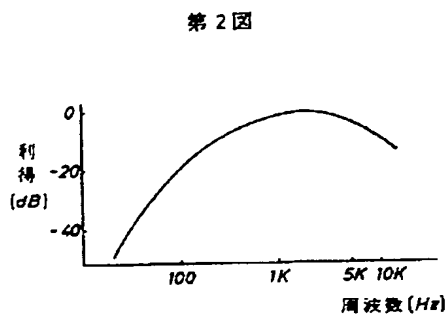
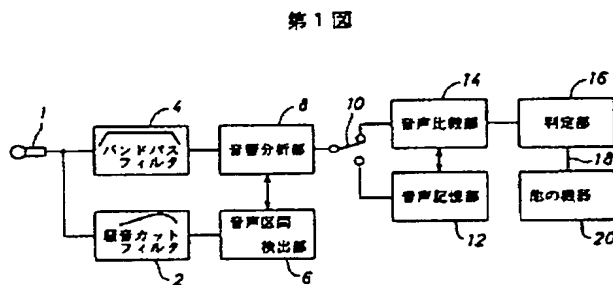
8…音響分析部

12…音声記憶部

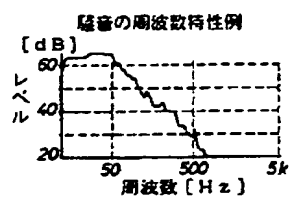
14…音声比較部

代理人 弁理士

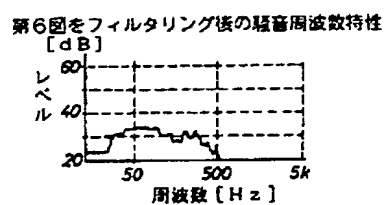
足立 勉(ほか2名)



第6図



第7図



第8図

